

A Practical Guide for Designing Future-Ready AI-Integrated Assessments in Education¹

Dr. Farhad Mehdipour, Director of Research and Head of Department

*Otago Polytechnic-Auckland International Campus (OPAIC), Auckland, New Zealand
farhadm@op.ac.nz*

Abstract – Artificial intelligence (AI) has revolutionized education, offering transformative opportunities across disciplines while presenting unique challenges for assessment design. In an era where AI tools are integral to learning, the question is not whether to use AI, but how to design assessments that ensure students meet learning objectives while leveraging AI's potential for enhanced learning experiences. This article provides practical guidelines for creating AI-safe assessments that balance automation and manual effort, evaluate technical proficiency, and promote ethical and transparent AI usage. It serves as a dynamic resource, evolving alongside AI's expanding role in education. Through case studies and actionable strategies, this article equips educators to prepare students for an AI-driven future workforce while maintaining academic integrity and encouraging critical engagement.

1 Introduction

AI has become an integral part of modern education, reshaping how students and educators interact with knowledge. As AI-driven technologies become more pervasive, their impact on pedagogy and assessment continues to grow. Adaptive learning platforms, for example, have demonstrated the ability to improve student outcomes by up to 30% through personalised instruction, providing customised pathways that accommodate diverse learning styles and needs (O'Neil, 2016; Melbourne Centre for the Study of Higher Education, 2023). Intelligent tutoring systems further enhance this experience by offering instant feedback and scaffolding, allowing students to progress at their own pace while addressing gaps in understanding. By automating routine tasks and providing insights into complex problems, AI enables personalised learning and innovative solutions. Recent advancements in tools like Cadmus (Cadmus, n.d.) and Gradescope (Gradescope, n.d.) have introduced scalable and transparent assessment methods, allowing educators to track student progress and ensure equitable learning experiences. These tools align with frameworks like the AI Assessment Scale, which categorises tasks into AI-assisted and manual components, enabling educators to balance automation with critical engagement (AISafeDesign, n.d.). Such innovations address scalability challenges in large classes while maintaining meaningful student-educator interactions.

AI's role in education extends beyond efficiency—it also facilitates innovation. Tools like generative AI models empower students to explore complex ideas, simulate scenarios, and engage in experiential learning previously limited by logistical constraints. These capabilities not only enrich the learning experience but also prepare students for real-world applications, making them better equipped for future challenges in an AI-driven world. Despite its potential, integrating AI into education presents unique challenges. Questions about over-reliance, academic integrity, and the preservation of critical thinking remain central to discussions about its use in assessments. This guide aims to address these challenges by offering a structured approach to leveraging AI in education responsibly.

¹ This article represents the first version of an evolving resource aimed at addressing the challenges and opportunities of AI-safe assessment design. I welcome your feedback and suggestions to improve and refine it for its second version. Please feel free to share your insights by contacting me at farhadm@op.ac.nz.

It provides actionable strategies for educators to design assessments that balance automation and manual effort, ensuring students develop both technical proficiency and conceptual understanding.

This guide presents a comprehensive framework for integrating AI into assessments, considering factors like class size, the distinction between theoretical and practical courses, and the varying needs of undergraduate and postgraduate students. By addressing these variables, the guide offers scalable, adaptable solutions that enrich education while enhancing critical thinking and ethical AI usage (UNESCO, 2023). Case studies in data analysis and software development are used to illustrate the application of these principles in specific contexts.

2 Balancing Opportunity and Challenge

AI tools have the potential to transform education across disciplines. For example, in STEM fields, AI-powered tools like MATLAB (MATLAB, n.d.) and TensorFlow (TensorFlow, n.d.) have streamlined the development and testing of complex algorithms, enabling students to focus more on interpretation rather than computation. In the humanities, platforms like Voyant Tools allow for deep analysis of texts, such as identifying recurring themes or tracing linguistic evolution in historical documents (Google AI Principles, n.d.). These advancements illustrate AI's capacity to enhance learning outcomes while demanding careful integration to maintain critical engagement. In STEM, tools like Python libraries automate data analysis by providing prebuilt functions for tasks such as data cleaning, visualisation, and statistical modelling, simplifying workflows for students and researchers. These libraries empower students to focus on interpreting results rather than getting bogged down in repetitive coding tasks. In the humanities, platforms like Voyant Tools (Voyant Tools, n.d.) enhance textual analysis by allowing users to explore word frequencies, themes, and patterns in large text corpora, facilitating deeper insights into literary and historical materials (Khan Academy, n.d.).

Recent innovations, such as Cadmus (Cadmus, n.d.) and Gradescope (Gradescope, n.d.), provide additional tools to address scalability and transparency challenges. Cadmus allows educators to monitor iterative workflows, ensuring students refine their work systematically while maintaining engagement. Gradescope automates routine grading tasks, freeing educators to focus on more complex interventions like personalised mentoring or live discussions.

However, reliance on AI can lead to superficial engagement, where students bypass critical thinking in favour of automated outputs (O'Neil, 2016). For instance, a student might use AI to generate a complete data analysis report without understanding the underlying methods or assumptions, which could lead to incorrect conclusions if the AI outputs are not critically evaluated. This lack of engagement hinders the development of analytical skills and the ability to independently verify results. Educators must design assessments that balance AI's efficiency with manual efforts to ensure meaningful learning. For instance, theoretical courses can use AI tools for textual analysis, encouraging students to critique outputs and refine interpretations manually. Practical courses might emphasise AI-assisted prototyping while requiring manual optimisation to develop problem-solving skills. Large classes can leverage AI for scalable feedback systems, while smaller cohorts benefit from personalised AI-supported mentoring. Each scenario demands customised strategies to maximise AI's potential while ensuring robust learning outcomes. Table 1 provides a summary of the opportunities and challenges associated with integrating AI into education, highlighting its transformative potential while addressing the complexities it introduces.

Table 1: Opportunities and challenges of AI in education

Opportunities	Challenges
<ul style="list-style-type: none"> • Personalised learning pathways • Real-time feedback through adaptive platforms • Automation of routine tasks (grading, data analysis) • Facilitation of critical engagement with complex ideas 	<ul style="list-style-type: none"> • Risk of over-reliance on AI • Threats to academic integrity • Potential biases in AI outputs • Balancing automation with foundational skill development

3 Challenges and Solutions

Integrating AI into education presents several challenges, including over-reliance on AI tools, ethical concerns, and scalability issues. For instance, students in large classes may heavily depend on automated feedback tools, which could hinder their ability to independently analyse and critique results. Similarly, ethical concerns arise when biases in AI algorithms inadvertently affect grading or recommendations, potentially disadvantaging certain groups. In smaller courses, managing equitable access to advanced AI tools might be difficult, creating disparities in learning outcomes. These challenges highlight the necessity for tailored solutions, such as incorporating manual assessment components and robust bias detection strategies, to ensure fairness and meaningful engagement. The introduction of AI into assessments requires educators to navigate a complex landscape where technology enhances learning but also raises significant questions about equity, ethics, and engagement. Below, we expand on these challenges:

- Over-reliance on AI Tools** – Students may become overly dependent on AI for problem-solving, bypassing critical thinking and manual skills development (McMinn, 2023). For instance, relying solely on AI-generated visualisations may lead to a superficial understanding of underlying trends or data anomalies. To counteract this, educators can design tasks that require manual analysis alongside AI-assisted outputs, ensuring students engage critically with their findings. Tasks might involve using AI for initial data visualisation but require students to interpret trends manually and identify potential outliers or anomalies.
- Ensuring Fairness and Equity** – AI tools often reflect biases inherent in their training data. This can lead to inequitable outcomes, particularly for marginalised groups. Educators must address these biases to ensure fairness in assessments (AISafeDesign, n.d.; O’Neil, 2016). Tools like IBM AI Fairness 360 (IBM AI Fairness 360, n.d.) can help identify and mitigate such biases, providing educators with resources to ensure AI outputs remain equitable and transparent.
- Plagiarism and Intellectual Honesty** – The use of AI for content generation poses challenges in distinguishing original student work from AI contributions. This jeopardises the credibility and fairness of assessments. Requiring students to explicitly acknowledge AI tools used, document their contributions, and provide detailed descriptions of their manual efforts enhances accountability and transparency.
- Scalability in Large Classes** – Managing the integration of AI in assessments becomes increasingly complex as class sizes grow. Ensuring that all students receive meaningful feedback and have equitable access to AI tools is a significant logistical challenge. Solutions include using platforms like Gradescope (Gradescope, n.d.) for automated grading and analytics dashboards in learning management systems like Moodle to monitor trends and identify struggling students. Structured peer reviews and reflective templates can further enhance scalability while maintaining engagement.
- Ethical Implications of AI Use** – Ethical concerns, such as data privacy, the use of proprietary AI models, and the societal impact of AI-driven decisions, need to be considered in the design of assessments.

(UNESCO, 2023; O’Neil, 2016). Educators should include reflective components that require students to evaluate the ethical implications of their AI use, such as bias in datasets or the societal consequences of AI-driven decisions.

Table 2 summarises these challenges and their corresponding solutions:

Table 2: Challenges and solutions of using AI in assessments

Challenge	Solution
Over-reliance on AI tools	Incorporate manual tasks and reflective components to ensure students actively engage with the material. For example, require students to manually interpret trends after using AI for data visualisation
Ensuring fairness and equity in AI-generated results	Teach bias detection and mitigation using platforms like IBM AI Fairness 360 and provide structured guidelines for verifying AI outputs
Plagiarism concerns	Require students to explicitly acknowledge AI contributions and differentiate them from their manual work, enabling transparency and accountability
Scalability in large classes	Use AI tools (e.g., Gradescope) for automated grading and analytics dashboards in platforms like Moodle to monitor performance trends. Structured templates for reflections and collaborative peer reviews can further enhance scalability
Ethical implications of AI use	Include components in assessments that require students to analyse the ethical implications of their AI use, such as bias in datasets or the societal impact.

4 Designing Assessments Across Academic Levels

Designing assessments that cater to the diverse needs of undergraduate and postgraduate students requires a nuanced understanding of their academic levels and learning objectives. Undergraduate courses focus on building foundational skills and introducing basic concepts, while postgraduate studies demand more specialised knowledge, critical analysis, and original research contributions.

Effective assessment design must therefore align with these differing expectations, providing opportunities for students to engage with AI tools in ways that enhance their learning and development. The following sections outline strategies for integrating AI into assessments at undergraduate and postgraduate levels.

Undergraduate courses focus on building foundational skills. Assessments should guide students in using AI for basic tasks while encouraging manual interventions to deepen understanding. For instance, students might use AI to clean datasets but must manually document trends and interpret findings. These tasks help students develop a balance between automation and critical analysis, enhancing both technical and conceptual understanding.

Postgraduate courses demand advanced analytical skills and original research. Assessments at this level can include tasks requiring customisation of AI tools, critical evaluation of methodologies, and exploration of ethical implications. For example, postgraduate students might modify AI algorithms for specific research applications, reflecting on their theoretical and practical implications. These assessments prepare students for leadership roles in their fields by encouraging innovation and ethical decision-making.

To address the diversity of academic levels and effectively adapt to varying needs, assessments should align with these levels to ensure both undergraduate and postgraduate students maximise their learning outcomes while preparing for real-world challenges in their respective fields. Staged or progressive assessments and reflective components can be employed as follows:

- Staged or progressive assessments – Divide assignments into iterative components to promote continuous learning. For example, students could submit proposals, interim reports, and final analyses to demonstrate the progression of their understanding and engagement with AI tools.

- **Reflective Elements** – Incorporate reflective components that require students to analyse their use of AI and assess their manual contributions. Questions such as "How did your manual interventions refine AI outputs?" or "What ethical considerations emerged in your use of AI?" deepen their engagement.

5 Theoretical vs. Practical Courses

Integrating AI into education requires a nuanced understanding of the differences between theoretical and practical disciplines. Each type of course presents unique opportunities and challenges when incorporating AI tools into the learning process. In theoretical courses, AI offers a means to explore abstract concepts and complex frameworks, allowing students to engage deeply with the material. Practical courses, on the other hand, benefit from AI's ability to streamline repetitive tasks, enabling students to focus on skill acquisition and application. However, in both contexts, it is crucial to ensure that AI is used as a complementary tool that enhances learning without replacing foundational skills or critical engagement.

5.1 Theoretical Courses

Theoretical disciplines benefit significantly from AI tools that facilitate deep engagement with complex concepts and frameworks. For example, in philosophy, students might analyse AI-generated summaries of ethical theories, comparing them to original texts to identify biases or omissions, thus sharpening their critical analysis skills. Similarly, in history, tools like Semantic Scholar can help students trace the evolution of historiographical debates, enabling them to explore patterns, shifts in interpretation, and the context of primary sources. These tasks encourage students to integrate AI outputs into their broader understanding of theoretical frameworks while facilitating analytical rigor. In addition, theoretical courses can leverage AI to model abstract concepts. For example, in economics, AI-driven simulations allow students to observe the dynamics of supply and demand in hypothetical markets. By critiquing the outputs of these simulations, students can test hypotheses and refine their understanding of underlying principles. Such activities illustrate the potential of AI to complement traditional theoretical learning methods while maintaining a focus on critical thinking and synthesis.

5.2 Practical Courses

Practical disciplines emphasise the acquisition and application of tangible skills. AI tools in these contexts serve as powerful assistants rather than replacements. For instance, in programming, students might use GitHub Copilot (GitHub Copilot, n.d.) for initial code generation but are required to debug and optimise the code manually to deepen their understanding of algorithms and best practices. Similarly, in architecture, AI can generate design prototypes, which students refine based on manual calculations, creative adjustments, and client requirements, ensuring that they retain essential problem-solving and design skills. To further illustrate, in engineering courses, AI-based tools like MATLAB can assist with circuit design or structural simulations. Students might start with an AI-generated model but are tasked with identifying its limitations and making necessary adjustments. This approach ensures that students develop a practical understanding of core concepts while benefiting from AI's efficiency. By balancing automation with manual refinement, practical courses prepare students to use AI effectively in professional settings, where precision and adaptability are paramount.

To maximise AI's potential across both theoretical and practical disciplines, educators can employ cross-disciplinary strategies that balance automation with critical engagement. Requiring manual refinement and interpretation of AI-generated outputs ensures that students actively engage with their tasks, regardless of the subject focus. For example, students might critique and improve AI-generated insights in a data analysis project, enhancing their understanding of underlying principles. Incorporating reflective components into assessments encourages students to critically evaluate AI outputs, documenting how their manual interventions have improved the results. Such

reflective exercises not only deepen their analytical skills but also build a thoughtful approach to integrating AI into academic work. Transparency is another vital aspect; clearly defining the role of AI in assignments ensures academic integrity and helps students understand the limitations of the tools they use. By integrating these strategies, educators can ensure that AI serves as a complement to, rather than a replacement for, foundational learning and skill development.

6 Scalability in AI-Assisted Assessments

The integration of AI into assessments must consider the scalability challenges and opportunities presented by different class sizes. In large classes, AI tools can alleviate workload pressures and improve consistency in evaluation, while in smaller classes, they can enhance personalisation and deepen student engagement (McMinn, 2023). Striking the right balance between AI-assisted automation and educator-led interventions is key to ensuring that assessments remain effective, equitable, and meaningful across varied learning environments (Melbourne Centre for the Study of Higher Education, 2023).

6.1 Large Classes

Scalability is crucial for large classes. AI tools (e.g., Gradescope) automate grading, providing consistent and timely feedback. Analytics dashboards in platforms like Moodle allow educators to monitor performance trends and identify areas where students struggle. These tools streamline the grading process, enabling instructors to focus on strategic interventions rather than routine tasks. For example, educators can use these analytics to identify common errors in student submissions and address them through targeted workshops or supplementary materials.

Despite these advantages, large classes pose challenges in maintaining engagement and equitable access to AI tools. Students may feel disconnected in environments that heavily rely on automated processes. To mitigate this, educators can incorporate live discussions, structured group activities, and opportunities for peer feedback. AI can support these initiatives by managing administrative tasks, such as organising breakout sessions or collating peer review results, allowing instructors to focus on encouraging meaningful interactions.

6.2 Small Classes

In smaller classes, AI supports personalised learning. Tools like Notion AI (Notion AI, n.d.) can create individualised study plans, while platforms like Otter.ai (Otter.ai, n.d.) assist students in summarising discussions and lectures. These tools free educators to focus on mentorship, creating a more interactive and tailored learning experience. For example, in a software development course, students might use AI to prototype solutions while engaging in iterative design discussions with their instructor, blending AI's efficiency with hands-on learning.

However, the flexibility of smaller classes allows for more experimental and creative uses of AI. Instructors can assign open-ended projects where students use AI to explore novel solutions to real-world problems. For instance, in a data science course, students could be tasked with designing a custom predictive model using AI, followed by a reflective analysis on its limitations and potential biases. This approach not only deepens their understanding of AI tools but also encourages critical thinking and innovation.

Small classes also face challenges, particularly in ensuring all students engage meaningfully with AI tools. To address this, educators can set clear guidelines on how AI should be used in assessments, requiring students to document their processes (AISafeDesign, n.d.) and critically evaluate the outcomes. This ensures accountability while encouraging a deeper appreciation for the technology's capabilities and limitations.

6.3 Addressing Engagement and Equity

To ensure meaningful engagement in both large and small classes, educators should specify clear guidelines for AI tool usage in assessments, requiring students to document their processes comprehensively. This approach promotes transparency and accountability, ensuring students critically engage with both the technology and their learning objectives. Reflective components should also be integrated into assessments, encouraging students to evaluate the limitations and ethical implications of their AI usage. Collaborative activities can be designed where students use AI tools to complement teamwork, encouraging the development of both interpersonal and technical skills. This balance of individual accountability and group interaction ensures a robust and meaningful learning experience.

7 Ethical and Transparent AI Use

The ethical integration of AI into education (UNESCO, 2023) ensures that its potential is harnessed responsibly while preserving academic integrity and critical thinking. As AI becomes increasingly prevalent, educators must guide students in understanding the ethical implications of their work and using AI tools with transparency and accountability (Melbourne Centre for the Study of Higher Education, 2023; McMinn, 2023). This involves embedding ethical considerations into the curriculum and providing students with clear frameworks for responsible AI usage.

7.1 Transparency

Students should document AI's role in their work, detailing the tools used, their contributions, and manual refinements. For instance, a data science project might include a workflow diagram illustrating which steps involved AI assistance and which were performed manually. This documentation ensures that students understand the scope and limitations of the tools they use while demonstrating their contributions to the work. Such transparency also ensures that students and educators alike can trace the thought processes behind a project, reinforcing the importance of personal accountability.

7.2 Ethical Considerations

Embedding ethical considerations into assessments encourages students to critically evaluate the broader societal implications of AI. Students should be guided to identify potential biases in AI outputs, assess the implications of these biases, and propose strategies to mitigate them. For instance, in a history project, students might reflect on whether an AI tool has amplified certain historical narratives or omitted others, prompting discussions about inclusivity and representation. Similarly, in STEM fields, students could examine how biases in training datasets might affect the reliability of AI models and their potential real-world applications. These reflective components enhance students' understanding of the ethical dimensions of AI use, preparing them to navigate related challenges in both academic and professional contexts.

7.3 Critical Thinking

Encouraging students to critically evaluate AI-generated insights is key to maintaining rigorous academic standards. Assessments should require students to compare AI-generated outputs with traditional methods, explore their limitations, and refine their understanding of the subject matter. For example, in an economics course, students could compare market analyses produced by AI tools with manual calculations, assessing the accuracy, assumptions, and relevance of each approach. Such exercises help students develop critical thinking skills while deepening their understanding of the subject.

7.4 Balanced Contributions

To ensure a comprehensive learning experience, tasks should involve a balance of AI-assisted and manual work. This approach ensures that students develop mastery of foundational skills while benefiting from the efficiency of AI tools. For instance, in a programming assignment, students might use AI for initial code suggestions but be required to debug, optimise, and annotate the final code manually. This combination of automation and manual effort enhances both technical proficiency and critical engagement. Similarly, in humanities projects, students could use AI to generate preliminary summaries but refine their arguments or analyses independently, demonstrating their ability to engage critically with AI-generated content.

8 Designing Assessments for AI Integration

The effective integration of AI into assessments requires careful planning and thoughtful design to ensure that both the benefits and challenges of AI are addressed. Educators must strike a balance between leveraging AI's capabilities and encouraging students' critical thinking and foundational skills. This section outlines a comprehensive framework for creating assessments that align with these objectives, providing practical strategies and examples to navigate this complex landscape effectively.

i) Clearly define learning objectives: Educators must articulate the skills and knowledge students are expected to demonstrate. Learning objectives should specify which aspects of the task can involve AI assistance and which require manual intervention to ensure balanced skill development. This clarity provides a roadmap for students to navigate AI-integrated tasks responsibly.

Example: A data analytics assignment might allow students to use AI tools to preprocess a dataset but require them to manually document the rationale behind data transformations or debug code generated by AI. Similarly, in a design course, students might use AI to create an initial prototype but must manually refine the design for usability and aesthetics.

ii) Include an AI usage plan: An AI usage plan facilitates structured engagement with AI tools. Students should outline how they intend to use AI in their projects, specifying the tasks for which AI will be used and where manual effort is required. This plan ensures transparency and alignment with the assessment's objectives.

Example: A student might state, "I will use GitHub Copilot for generating boilerplate code but will manually optimise it for security and performance. The final design will be reviewed to ensure compliance with project specifications."

iii) Emphasise documentation and transparency: Transparency in AI usage is critical for maintaining academic integrity and ensuring accountability. Students should document:

- The specific AI tools used,
- Tasks completed using AI, and
- Manual contributions and decision-making processes.

This documentation serves as a reflective tool, allowing students to evaluate their engagement with AI and educators to assess the authenticity of their work.

Example: A workflow diagram or annotation that differentiates between AI-assisted and manual processes can highlight students' understanding and contributions.

iv) Incorporate reflective components: Reflective questions encourage students to critically evaluate their use of AI, particularly in programming projects. Prompts can include, "Why did you choose a specific algorithm suggested by an AI tool?" or "How did you manually improve the AI-assisted code to optimise performance and readability?" Such targeted reflections facilitate a deeper understanding of both AI

capabilities and the student's own contributions. Questions such as, "What limitations did you encounter with the AI tool, and how did you address them?" encourage deeper engagement.

Example: A student reflecting on AI-generated code might note, "The generated algorithm lacked error handling, which I implemented manually to ensure robustness."

v) Develop rubrics to evaluate AI integration: Rubrics are essential for providing clear and structured assessment criteria, especially when integrating AI into educational tasks. To evaluate AI integration effectively, rubrics should encompass the following dimensions:

- a. Understanding and application of subject-specific skills – This criterion evaluates students' ability to apply the skills and frameworks taught in their discipline effectively. As it forms a fundamental part of any course rubric, it ensures students demonstrate relevant expertise, whether through technical, analytical, or conceptual methods. While AI may augment learning, this criterion remains universally applicable to assess students' mastery of course-specific objectives.

Example: A history student might analyse the use of AI in interpreting primary sources, comparing AI-generated summaries with their manual analyses.

- b. Critical evaluation of AI outputs or code – Examine students' ability to critique the quality and reliability of AI-generated outputs or code. This involves assessing their identification of biases, errors, or limitations and their proposed solutions to address these issues.

Example: Did the student identify and correct biases in an AI-generated prediction model?

- c. Transparency in documenting workflows – Ensure students clearly document their AI usage, including which steps involved AI and which were completed manually. Transparency highlights their engagement with both automated and manual processes.

Example: Did the student provide a clear visualisation of their workflow?

- d. Ethical considerations – Evaluate students' awareness and application of ethical principles in AI usage. This includes addressing data privacy, algorithmic bias, and the societal impact of their AI applications.

Example: Did the student mitigate ethical concerns in using sensitive datasets?

- e. Originality and manual contributions – Assess the balance between AI-assisted work and manual effort. Students should demonstrate independent analytical skills and creativity beyond AI's capabilities.

Example: Did the student enhance AI-generated code with original logic or design a novel solution based on AI recommendations?

To synthesise the core concepts discussed in this article, Fig. 1 and Table 2 provide a structured overview of the guidelines and their application across different course types. These summaries serve as quick references for educators aiming to implement AI-safe assessments effectively. They highlight the key steps in assessment design and illustrate how these principles can be tailored to theoretical and practical courses.

Fig. 1: Key steps for designing AI-safe assessments

Define Learning Objectives	Plan AI Usage	Incorporate Ethical Guidelines	Align Assessments by Course Type and Level	Develop Rubrics	Embed Reflective Components	Transparency in Documentation	Deliver Assessment	Evaluate and Refine
<ul style="list-style-type: none"> Set clear goals for the skills and knowledge students should demonstrate. Example: Require students to use AI for generating initial outputs but rely on manual analysis for refinement. 	<ul style="list-style-type: none"> Develop a plan outlining where AI tools will be used and where manual work is required. Example: Use AI to summarize data but interpret insights manually. 	<ul style="list-style-type: none"> Ensure students understand ethical AI usage, including addressing biases and privacy concerns. Example: Reflect on ethical challenges such as biases in AI-generated suggestions. 	<ul style="list-style-type: none"> Adjust the tasks depending on whether the course is theoretical or practical, undergraduate or postgraduate. Example: Practical tasks may include debugging AI-generated code, while theoretical tasks might involve critiquing AI-summarized theories. 	<ul style="list-style-type: none"> Create clear evaluation criteria that assess both AI integration and manual contributions. Example: Include transparency, critical thinking, and ethical awareness in the grading rubric. 	<ul style="list-style-type: none"> Encourage students to document and reflect on their use of AI. Example: Reflect on why a specific AI-generated suggestion was accepted or modified. 	<ul style="list-style-type: none"> Require students to document their AI usage and manual refinements. Example: Submit a workflow diagram that clearly differentiates AI-assisted and manual tasks. 	<ul style="list-style-type: none"> Students complete and submit their work, including all required documentation and reflections. 	<ul style="list-style-type: none"> Educators review the assessments, provide feedback, and refine guidelines for future tasks.

Table 2: Applying AI-safe assessment principles across theoretical and practical courses

Step	Theoretical Courses	Practical Courses
Define learning objectives	Focus on conceptual understanding and critical analysis.	Focus on skill development and application.
Plan AI usage	Use AI for summarisation; refine insights through manual analysis.	Use AI for prototyping or initial outputs; manually debug or optimise.
Incorporate ethical guidelines	Explore biases in AI-generated theoretical frameworks.	Address real-world ethical implications, such as data privacy and bias.
Tailor assessments by course type and level	Emphasise theoretical engagement and critical debates.	Balance automation with manual skill application (e.g., debugging).
Develop rubrics	Evaluate analytical depth, understanding of theories, and critique quality.	Assess output quality, manual interventions, and skill proficiency.
Embed reflective components	Encourage reflection on AI outputs and their alignment with theoretical frameworks.	Reflect on problem-solving processes and AI's role in task execution.
Transparency in documentation	Document rationale for AI-assisted conclusions and manual analyses.	Document AI tool usage and manual improvements to outputs.
Deliver assessment	Explain theoretical frameworks, limitations, and challenges.	Provide functional solutions supported by clear documentation.
Evaluate and refine	Use feedback to refine theoretical frameworks and improve clarity.	Iterate on practical solutions and address real-world limitations.
Case studies	Provide examples where students critique AI-generated concepts or models.	Showcase tasks like debugging AI-generated code or refining prototypes.
Impact	Demonstrate how AI aids in deep conceptual understanding and analysis.	Highlight how AI supports skill-building and real-world application.

9 Guiding Students in AI Use

Integrating AI into educational contexts requires equipping students with the knowledge and skills needed to use these tools effectively and responsibly. Educators must offer clear guidance on the capabilities and limitations of AI, balancing the benefits of automation with the development of independent analytical and problem-solving skills. This section explores strategies for teaching students to critically engage with AI, adhere to ethical standards, and apply their expertise effectively. By embedding these strategies into the curriculum, educators can support students in navigating the complexities of AI integration, ensuring its responsible and impactful use. These approaches

nurture critical thinking, ethical awareness, and technical proficiency, preparing students for success in an AI-driven world.

- Teaching AI tool basics – Educators should familiarise students with commonly used AI tools, highlighting both their technical capabilities and inherent limitations. Offering workshops or tutorials can effectively address knowledge gaps and build foundational skills.
- Encouraging critical thinking – Students should be taught to question AI outputs. For instance, if a regression model predicts a negative sales value or if an AI tool generates inefficient code, students must recognise these as errors and investigate their causes.
- Ethical AI usage – Ethical considerations must be integral to teaching. Students should understand data privacy, avoid plagiarism, and ensure that AI does not perpetuate biases or shortcuts in programming. For example, students using AI for code generation should ensure the generated code adheres to copyright and licensing requirements.
- Balancing automation and manual work – While AI can handle repetitive tasks, students must demonstrate their analytical and coding capabilities. For example, a student might use AI to identify trends in a dataset or generate a code snippet but manually interpret implications or optimise the code for performance.
- Scenario-based learning – Introduce real-world scenarios where students decide whether and how to use AI. This helps contextualise their learning and promotes decision-making skills. For instance, in an ethics course, students could evaluate the implications of using biased AI algorithms in hiring processes, exploring potential societal impacts.
- Feedback and iteration – Provide structured feedback on AI usage plans and outputs. This iterative process reinforces best practices and encourages improvement over time. For example, students submit a preliminary AI-assisted project plan, receive feedback on areas needing transparency or ethical consideration, and revise accordingly.
- Integrating Peer Review – Students can review each other's AI-assisted outputs to learn from diverse perspectives and collaboratively identify strengths and weaknesses. For example, in an information systems course, students evaluate peer-developed AI-generated workflows for robustness and adherence to ethical standards.

10 Case Studies

The case studies in this section offer practical insights into the integration of AI in educational assessments, demonstrating how it can enhance learning outcomes while preserving academic integrity and enabling critical engagement. These examples illustrate AI's application in both postgraduate and undergraduate courses, highlighting strategies to balance AI assistance with manual effort and address ethical considerations.

Case Study 1: Integrating AI in a postgraduate Data Science course

In a postgraduate data science course, students are tasked with predicting customer churn using AI models. The assessment included the following components:

- i) AI Usage Plan: Students outline the tasks for AI tools, such as data preprocessing or feature selection, and specified areas requiring manual analysis to ensure transparency and accountability.
- ii) Critical Reflection: Reflective reports require students to evaluate model performance, identifying biases or inaccuracies in AI predictions, such as overfitting or poor handling of outliers. Students also explore the implications of these findings on decision-making processes.

- iii) Customised Models: Students customise AI algorithms, such as modifying hyperparameters or incorporating additional features, to enhance prediction accuracy and demonstrate advanced analytical skills.

This assessment highlights the dynamic interplay between AI assistance and student expertise. It enables critical engagement with AI-generated insights while promoting ethical awareness and advanced technical skills. The project's structure ensures that students not only enhance their analytical abilities but also develop the capacity to critically evaluate and refine AI outputs effectively.

Case Study 2: AI-safe assessments in an undergraduate Information Systems course

In an undergraduate information systems course, students are tasked with designing a conceptual framework for implementing a cloud-based solution for a small business. The assessment includes the following components:

- i) AI usage plan: Students utilise AI tools such as ChatGPT (ChatGPT, n.d.) or Google Bard (Google Bard, n.d.) to generate initial ideas for the framework, including identifying potential benefits and risks of cloud integration. However, they are required to critically evaluate and refine these AI-generated ideas using academic literature, industry case studies, and their analyses to develop a robust and informed conceptual model.
- ii) Critical reflection: Reflective components are integrated into the assessment, requiring students to identify the limitations of AI-generated suggestions. This included addressing gaps in the AI outputs, recognising potential biases in the information provided, and questioning assumptions inherent in AI recommendations. Students are encouraged to document how these limitations influenced their final framework.
- iii) Manual enhancements: Students manually create diagrams or workflow models to articulate their cloud-based solution. These models demonstrate their understanding of the business requirements and provide evidence of their ability to integrate practical implications with theoretical knowledge. For example, students design process maps and data flow diagrams that illustrated the operational flow of the proposed system.

This assessment allows students to engage critically with AI-generated content, combining automated insights with manual synthesis and validation. The project's structure ensures that students develop a foundational understanding of theoretical and practical aspects of cloud-based information systems, while also addressing ethical considerations of AI use in professional settings.

9 Conclusion

The integration of AI into education represents a transformative opportunity to enhance learning experiences across disciplines. However, realising its full potential requires careful and thoughtful design to address the challenges it presents. By adapting assessments to academic levels, course types, and class sizes, educators can create meaningful opportunities for students to leverage AI responsibly while encouraging critical engagement with the material. Transparent and ethical AI usage is central to maintaining academic integrity and developing essential skills such as critical thinking, ethical awareness, and technical proficiency. Through balanced and well-structured assessments, educators can encourage students to explore the capabilities of AI tools, critique their limitations, and apply their learning in diverse, real-world contexts. This approach not only redefines traditional education but also prepares learners to thrive in an increasingly AI-driven world. By embracing innovation while maintaining a commitment to foundational skills and ethical principles, educators can empower students to address emerging challenges with confidence, adaptability, and responsibility.

References

- 1) AISafeDesign. (n.d.). What is AI safe?. Retrieved from <https://aisafedesign.com/what-is-ai-safe/>
- 2) Cadmus. (n.d.). Process-oriented assessment tools for education. Retrieved from <https://www.cadmus.io/>
- 3) ChatGPT. (n.d.). AI-powered conversational agent. Retrieved from <https://openai.com/chatgpt>
- 4) GitHub Copilot. (n.d.). Your AI pair programmer. Retrieved from <https://github.com/features/copilot>
- 5) Google AI Principles. (n.d.). Retrieved from <https://ai.google/principles>
- 6) Google Bard. (n.d.). Experimental conversational AI service. Retrieved from <https://bard.google.com>
- 7) Gradescope. (n.d.). AI-assisted grading platform. Retrieved from <https://www.gradescope.com>
- 8) IBM AI Fairness 360. (n.d.). Retrieved from <https://aif360.mybluemix.net/>
- 9) Khan Academy. (n.d.). Personalized learning. Retrieved from <https://www.khanacademy.org>
- 10) MATLAB. (n.d.). MATLAB software for engineering and science. Retrieved from <https://www.mathworks.com/products/matlab.html>
- 11) McMinn, S. (2023). Incorporating AI into learning and assessments: A guided pathway. Retrieved from <https://www.linkedin.com/pulse/incorporating-ai-learning-assessments-guided-pathway-sean-mcminn/>
- 12) Melbourne Centre for the Study of Higher Education. (2023). AI and assessment: Guidelines for educators. Retrieved from <https://melbourne-cshe.unimelb.edu.au/resources>
- 13) Notion AI. (n.d.). Your AI-powered workspace. Retrieved from <https://www.notion.so/product/ai>
- 14) O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.
- 15) Otter.ai. (n.d.). AI transcription and note-taking tool. Retrieved from <https://otter.ai>
- 16) TensorFlow. (n.d.). An end-to-end open source machine learning platform. Retrieved from <https://www.tensorflow.org>
- 17) UNESCO. (2023). Ethical guidelines for AI in education. Retrieved from <https://unesco.org/ai-guidelines>
- 18) Voyant Tools. (n.d.). Text analysis for digital humanities. Retrieved from <https://voyant-tools.org>.